MIRS APPLICATIONS

MIRS (**M**aster **I**nformation **R**etrieval **S**ystem) is a proprietary software system for people who have, or are developing, collections of information. MIRS handles the work of turning your information into an organized data base. From a MIRS data base you can instantly retrieve information, answer questions, produce selective reports, and update data.

For inventory control, medical diagnosis, marketing studies, investment analysis, scientific data retrieval, reservations systems – in short for any information retrieval application which **requires immediate system response**, MIRS offers a uniquely powerful and flexible capability.

MIRS BENEFITS

MIRS is flexible. MIRS has virtually no format or size restrictions. It is designed so that you can use it as a simple tool, or as the inner works of a custom–tailored system. You can specify your own formats, and your own terminology – and you can use your existing data files.

MIRS is economical. MIRS storage automatically expands and contracts as the user expands and contracts his data base. MIRS modular design and full range of information retrieval techniques mean that it can adapt itself to the characteristics of a particular application and provide the most cost–effective solution for that application.

MIRS is easy to use. MIRS uses concise and simple terms to establish a dialog with the user. Anyone who can get a book out of the library can be taught to use MIRS in one practice session.

MIRS simultaneously provides large numbers of users with instant access to the same data base. Each user can update the data base as well as retrieve information from it. The effects of an update by one user are immediately available to all other users.

HOW MIRS WORKS

MIRS is analogous to the information retrieval system used by libraries, in which the books, listed in a master card catalog, are numbered and assigned one or more "descriptors", such as title, author, and subject(s). Cards in a library catalog are arranged alphabetically by descriptor– name so that the user may look up any descriptor and find the catalog numbers of all books to which it applies. In such a system there are two broad classes of functions:

1. **Query/Retrieval Functions** with which a user specifies the books he wants and retrieves them, without altering the catalog or library structure; and

2. **Librarian Functions** in which new books are introduced and catalogued, obsolete volumes and their references removed, and alterations to the catalog made without changing the book collection itself.

MIRS may be thought of as a time-sharing version of the library system described above, in which "data item" replaces "book" as the basic unit of information stored, catalogued and retrieved. Data items may be anyting – – an invoice, the text of War and Peace, or a chemical formula.

The way you express your inquiry tells MIRS what to look for. The simplest query is one word such as:

ENGINEER

Most queries include expressions that more precisely specify which items are of interest. They might designate whether the person you want is: an ENGINEER **and** MARRIED, an ENGINEER **or** MARRIED **or both**, an ENGINEER **and not** MARRIED, an ENGINEER **or** MARRIED **but not both**.

The circles in the diagrams below represent all the engineers (E) and all the married persons (M) in a hypothetical data base. The shaded areas correspond to the persons designated by each query.


ENGINEER
and
MARRIED
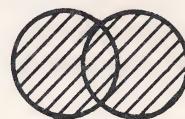
ENGINEER
or
MARRIED
or both
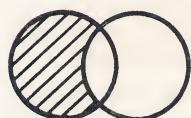
ENGINEER
and not
MARRIED
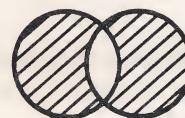
ENGINEER
or
MARRIED
but not both

The MIRS system is a powerful tool. With it the user may cross-reference and compile a list of data items which satisfy an arbitrarily complex set of criteria.

Many users may simultaneously interrogate a MIRS catalog or data base, even while updating is in progress. Unlike other similar systems, whose data bases must be updated periodically in batch mode, MIRS permits total interactive usage and updating at any time.

DATA BASE SECURITY
Access to the data bases created by MIRS may be restricted to selected groups of users. AL/COM LOGIN codes, which users may change periodically, offer three "password" levels of security.

MIRS AVAILABILITY
MIRS is ready to work for you today. Your AL/COM representative can show you MIRS in action with sample data bases. He can give advice and assistance in setting up a data base. For data already in existence on cards, tapes, etc. he can help you use a wide variety of methods to easily get your data into a MIRS data base.

CREATING A MIRS DATA BASE

Items may be entered into a MIRS data base at any time, even while others are using the same data base, by typing:

ENTER–ITEM–TEXT–AND–DESCRIPTORS

MIRS responds by asking for the body of the item you are entering:

TYPE TEXT:

You would type:

JOHNSON, GERALD F.
SWITCH DESIGN, CALIFORNIA :::

(Typing "::: " signals the end of the body of the item)

MIRS then asks for the "descriptors" under which you want the item cataloged:

TYPE DESCRIPTORS:

and you give them:

BSEE MARRIED.

MIRS immediately puts the new item in the data base.

For greater efficiency, or to handle data base items prepared by other computer programs, MIRS also allows the entry of batches of items from computer files. Updating of existing items is also allowed.

These are the basic inquiry and entry functions of MIRS. In addition, there are various service functions available. There are many other ways you can handle information with MIRS, always with simple English words.

MIRS is efficient for both large and small applications. If desired, special features or custom–tailored versions of MIRS can be easily implemented from its modular design. Your AL/COM representative will help you analyze your application.

USING A MIRS DATA BASE

To use a MIRS data base all that is necessary is to dial the AL/COM telephone number, identify yourself with your special log-in codes, and indicate that you wish to use MIRS with your data base. Simply type:

RUN MIRS DB

Specifying the items you want to retrieve is easy. MIRS uses familiar words and phrases put together in a natural way. For example, suppose your data base contains information about personnel in your company and you need to find an engineer, either electrical or mechanical, for a particular assignment. You might type into MIRS:

QUERY BSEE OR BSME.

MIRS would respond with the number of candidates:

75 HITS. WHAT NOW?

Not wanting to look at full information on that many candidates, since it's an assignment calling for a lot of traveling, you might decide to look first at the unmarried men. You would type:

QUERY-EXTEND AND-NOT MARRIED.

and finding that:

21 HITS. WHAT NOW?

you might decide to have full information on all of them typed out:

RETRIEVE-TYPE

MIRS would then type out all the information from the data base on these 21 candidates.

MIRS has a number of facilities for using the results of an inquiry. Among them are having the items typed out in full (RETRIEVE-TYPE), having the items put in computer storage for processing by other programs (RETRIEVE-FILE), and having them turned into a simple report (REPORT).

Inquiries can be complex if you wish. Just use parentheses (as in arithmetic and algebra) to make clear how the and's, or's, etc. go together. For example:

QUERY (PROGRAMMER OR ((BSEE OR BSME) BUT-NOT PHD)) AND-NOT MARRIED.

This is equivalent to asking MIRS for "either programmers or engineers, with BSEE or BSME, but not with a doctorate degree, who are not married".

MIRS is efficient because it can answer questions without having to search through an entire data base; MIRS catalogs the items for easy reference, as in a library. In the example above, MIRS has lists of the BSEE's and BSME's and can quickly count the list of those who are either BSEE or BSME without having to examine all personnel items to see who is a BSEE or BSME. MIRS also has the ability to search through the data "sifting" items to select particular ones. Suppose you have a bibliography but haven't bothered to have MIRS catalog items by publication date or by the words and parts of words in titles. You could, for example, select just those items published before 1964 and having "auto" somewhere in the title by typing:

SIFT PUB-DATE LESS-THAN 1964 AND TITLE CONTAINS "AUTO".

Sifting through items is simple, but if you wish, complex sifting can be done. Arithmetic can be used, as in:

SIFT STOCK-ON-HAND TIMES UNIT-COST GREATER-THAN 100 OR
MINIMUM-ORDER TIMES UNIT-COST GREATER-THAN 100.

QUERY, QUERY-EXTEND, and SIFT can be used over and over, in any combination, to refine or widen a search in order to find exactly the items you want.

For convenience in typing, you may abbreviate. If you are mathematically oriented, you may use a more algebraic notation. For example, the following all mean the same thing to MIRS:

QUERY-EXTEND (BSEE OR BSME) BUT-NOT PHD.
QE (BSEE, BSME) BN PHD.
QE (BSEE + BSME) >-PHD.

MIRS (**M**aster **I**nformation **R**etrieval **S**ystem) is a proprietary software system which enables multiple users to simultaneously create, interrogate, and update central data bases in a time-sharing environment.

It is available exclusively on the AL/COM nationwide time-sharing network, a service of Applied Logic Corporation.

This document gives a technical overview of MIRS and MIRS data bases. It complements the "MIRS Introductory Information", which contains an introduction to the features of MIRS and has examples.

MIRS DATA BASE STRUCTURE

A MIRS **data base** consists of a set of items and a catalog containing retrieval information for the items.

The **catalog organization** is that of inverted lists for user chosen descriptors applying to the various items. Retrieval operations include use of the inverted lists and other techniques including sifting through the body of items. The catalog is basically a hash table indexed by hash codes for the descriptors. A method called the "cluster buster" technique using small in-core tables ensures that the page corresponding to a descriptor can be obtained in exactly one disk read operation regardless of clustering and filling up of pages.

Short **inverted lists** are stored with descriptors in the catalog. Long inverted lists are stored outside the catalog.

Data base items are stored as variable length records packed together so that no space is wasted. The user may choose (independently for each item-descriptor pair) whether to store and retrieve the descriptor with the body of the item.

There are **no size limitations**, up to the capacity of physical mass storage, on the size of the data base, number of items, length of individual items, size of the catalog, length of inverted lists, etc. Each AL/COM time-sharing system, a dual PDP-10 configuration, has 500,000,000 characters of on-line storage. The splitting of logical files into AL/COM system files and the expansion of item files, the catalog, hash code length, etc. is handled automatically. Contraction to recover space is an available service function.

MIRS USER INTERFACE

English words and hyphenated phrases may be used for all vocabulary. Short forms and algebraic notation for operators may be intermixed with long forms.

Descriptors are usually strings of characters containing alphanumerics and dashes. Descriptors containing other characters are allowed when enclosed in quotation marks. The current limit (easily expanded) is 120 characters per descriptor.

Help is available on-line. Response to a "?" from the user in any situation is an explanation of options available, meaning of error message, etc. as appropriate. Repeated "?" will yield fuller explanations.

The user interface is structured so that the standard MIRS dialog and action can be replaced in whole or in part by custom-tailored subroutines. Input and output from MIRS can be translated and modified as desired.

MIRS INQUIRIES, RETRIEVALS, UPDATING

Basic **queries** are Boolean expressions containing descriptors. The **hitlist** resulting from a query may be further refined or expanded by a **query extension** which takes the form of a Boolean operator and expression. Conditions involving the content of items may be used in obtaining hitlists. Such **"sifting"** is specified by logical expressions containing names of fields within the item, numeric and string constants, arithmetic operators, arithmetic relations, elementary string relations, and Boolean connectives. The sets of arithmetic operators, arithmetic relations, and Boolean connectives are complete. There is no practical limit on length, depth, or nesting in arithmetic and Boolean expressions. Query, query-extension, and sifting may be repeated indefinitely and in any order. Refinement and expansion of hitlists may be interspersed with retrieval operations on the hitlist items.

Retrieval options include typeout on the terminal, browsing, writing of a file for later use, and generation of simple reports.

Updating may be done interactively or from prepared files. Updating includes deleting items, modifying items, adding items, removing descriptors from the catalog, and defining new descriptors for existing items.

Multiple simultaneous inquiries and updates are allowed. A batch mode of updating, which gains efficiency at the price of locking out other users, is also available.

MIRS AND GENERAL AL/COM FEATURES

The AL/COM Network makes possible nationwide access to a common MIRS data base.

AL/COM login and file security procedures allow selective control over inquiry and update access to a data base.

User programming for processing MIRS input and output files may be done in FORTRAN, COBOL, SNOBOL, LISP, and assembly languages. MIRS itself is programmed in a combination of assembly and higher level languages.

Back-up protection for program and data files is provided automatically by the AL/COM file back-up system. Back-up copies are made on tape incrementally (usually within 15 minutes of modification under current scheduling) and additional back-up copies of files are made on a daily and weekly basis. Back-up protection for system failure is provided by multiple systems in the central AL/COM computer center in Princeton, New Jersey.